
CALCULATING THE STANDARD ERROR:
JUST HOW MUCH SHOULD EMPIRICAL STUDIES CURB
OUR ENTHUSIASM FOR LEGAL STANDARDS?

*Brian Sheppard**

Professor Shiffrin makes a novel case in favor of legal standards: they are better than legal rules at compelling moral deliberation and thereby improve the function of our democratic institutions.¹ While there is a growing chorus of scholars singing the praises of standards, Shiffrin's claim is unique because it takes the plain directives of standards very seriously. Standards direct us to apply moral terminology, so they make it more likely that we will engage in moral deliberation. In this way, her analysis of the positive effects of standards is simpler and more direct than others. By contrast, other scholars have liked standards because their vagueness can operate as a cost-efficient delegation of the labor of legal interpretation to law-applying organs,² they have a rhetorical value apart from their content,³ they allow for more transparent policymaking,⁴ they limit unfavorable legal gamesmanship,⁵ or they are simply the best way to capture contested or essentially vague concepts.⁶ More than these other scholars, Shiffrin reminds

* Associate Professor, Seton Hall University School of Law. I would like to thank Andrew Vahid Moshirnia for his valuable research assistance. All errors are my own.

¹ See Seana Valentine Shiffrin, *Inducing Moral Deliberation: On the Occasional Virtues of Fog*, 123 HARV. L. REV. 1214, 1223–25 (2010).

² See Louis Kaplow, *Rule Versus Standards: An Economic Analysis*, 42 DUKE L.J. 557, 621–22 (1992) (“If behavior subject to the law is infrequent, however, standards are likely to be preferable. Of particular relevance are laws for which behavior varies greatly, so that most relevant scenarios are unlikely ever to occur. Determining the appropriate content of the law for all such contingencies would be expensive, and most of the expense would be wasted. It would be preferable to wait until particular circumstances arise.”).

³ See Carol M. Rose, *Crystals and Mud in Property Law*, 40 STAN. L. REV. 577, 610 (1988) (“But it is precisely as metaphor or rhetoric that the choice between crystal and mud matters. . . . Thus, crystal rhetoric suggests that we view friends, family, and fellow citizens from the same cool distance as those we don't know at all; while mud rhetoric suggests that we treat even those to whom we have no real connection with the kind of engagement that we normally reserve for friends and partners. And for this reason — for the sake of the different social didactics, the different modes of conversation and interaction implicit in the two rhetorical styles — we debate endlessly the respective merits of crystals and mud.” (footnotes omitted)).

⁴ See Kathleen M. Sullivan, *The Supreme Court, 1991 Term — Foreword: The Justices of Rules and Standards*, 106 HARV. L. REV. 22, 67 (1992) (“On this view, standards make visible and accountable the inevitable weighing process that rules obscure.”).

⁵ See John C. Coffee, Jr. & Hillary A. Sale, *Redesigning the SEC: Does the Treasury Have a Better Idea?*, 95 VA. L. REV. 707, 751 (2009) (“A rule, because it is certain, does not allow for flexibility or substantive equality. It can be over- or under-inclusive, and can encourage behavior that is socially irresponsible up to the line it draws.” (footnote omitted)).

⁶ See Marc R. Poirier, *The Virtue of Vagueness in Takings Doctrine*, 24 CARDOZO L. REV. 93, 190 (2002) (“As an essentially contested concept, it is fertile and generative precisely because it is

us that the content of a legal standard often directs us to consider a very particular thing: morality. These moral standards call us to “consider what is or is not fair or reasonable or unconscionable — not merely what is or is not in [our] interest” such that we are involved “more directly in understanding and implementing the law” and stimulating “morally sensitive interaction between us.”⁷ It is a claim so essential that it has been overlooked or forgotten.

In addition, Shiffrin’s essay is a welcome contribution to the fields of conceptual and normative legal philosophy. As to the former, she identifies an intuitive relationship between standards and moral directives⁸ and then describes the conceptual contours of the moral deliberation that can arise from standards.⁹ As to the latter, she explains how inducing moral deliberation fosters our democratic ideals¹⁰ without running afoul of our interest in personal autonomy or fairness.¹¹

At bottom, however, this work is most valuable insofar as it is understood to be making an empirical claim — namely, that legal standards are more likely to induce moral deliberation, which, in turn, brings about moral behavior,¹² than are legal rules. If this is true, the notion that standards are second-rate is considerably weakened. Nev-

inevitably, and perhaps quintessentially, vague and unresolvable. It does not and cannot give clear rules at the level of generality and simplicity demanded of it.”)

⁷ Shiffrin, *supra* note 1, at 1231.

⁸ *Id.* at 1217–18, 1222–23.

⁹ *Id.* at 1226–31

¹⁰ *Id.* at 1223–26

¹¹ *Id.* at 1231–45

¹² While Shiffrin’s focus is moral deliberation rather than moral conduct, it is clear that she links an increase in both to the use of legal standards. She states, “one virtue of standards is that their lack of precision induces moral deliberation as well as the deployment and exercise of *moral skills*.” *Id.* at 1222 (emphasis added). She does concede, however, that the change will take time, stating, “[m]y celebration of induced deliberation does not imagine an unrealistic model in which individuals identify an evaluative term, come to a considered judgment about its proper application, and adjust their conduct accordingly on the spot. . . . The process by which deliberation affects behavior and articulates cognition may be more akin to a slow, sometimes clogged, drip than to a quick, direct injection.” *Id.* at 1226.

Even if it is not her primary focus, the conduct component is surely important. It could reasonably be argued that the moral deliberation that legal standards ask or command us to perform is valuable only to the extent that it makes those engaging in it behave more morally. In other words, challenges to the intrinsic value of moral deliberation cannot be dismissed out of hand. For my part, I am willing to admit that moral deliberation has intrinsic value, I just do not think that its value surpasses the value of considerably increasing actual moral conduct. For example, I can agree that the decision to help a person in need is more valuable if it is based on careful moral deliberation about the way to live a proper life than, say, on reflection about how to minimize tax liability. In contrast, if it could be shown that moral deliberation actually brought about even a slight, nontrivial overall *reduction* of moral conduct, then I would advocate that we not engage in it under those circumstances. Further analysis of this issue is beyond the bounds of this Essay. Suffice it to say here, I hope, that Shiffrin’s argument at the very least requires her to consider both whether standards bring about moral deliberation *and* whether they bring about moral conduct when she evaluates their worth.

ertheless, such a claim must answer to empirical research on the subject.

For her part, Shiffrin does not cite very much empirical evidence in support of her hypothesis.¹³ This omission is somewhat excusable; there simply have not been many published empirical studies of the comparative effect of rules and standards on those subject to them. Still, there has been some research, including my own, and the results do not support her empirical proposition.

Here, I first complicate the frame through which Shiffrin has chosen to analyze rules and standards to make it more reflective of the form and context of legal norms. Through this new, more complicated frame, I then review the empirical scholarship and suggest that Shiffrin scale back and fine-tune her claim about legal standards to make it more consistent with our knowledge of how people respond to them.

DISCUSSION

Broadly speaking, Shiffrin's essay is about the best way to draft legal norms for a particular purpose. Legal standards are a particular kind of norm or, more particularly, norm statement. In singling out legal standards, Shiffrin has drawn two types of distinctions among norm statements: one of norm type — specifically, between legal and nonlegal norms — and one of norm-content — specifically between rules and standards.¹⁴ She pays the most attention to the latter distinction, but the former distinction is important too.

We can identify all sorts of norm types, some more familiar or prominent than others. Many norm typologies distinguish norms based upon their source or context: for example, they list social norms, religious norms, etiquette norms, fashion norms, ideological norms, and, of course, legal norms as separate types.¹⁵ An important lesson to draw from these lists is that legal norms do not operate in isolation. Most of us are conscious that we are simultaneously subject to legal norms and nonlegal norms. We know that, right now, we are under legal, moral, and perhaps even religious norms that prohibit theft or murder, for example. In considering Shiffrin's claim, we ought to pay attention to nonlegal norms insofar as their interaction with legal norms either increases or decreases the amount of moral deliberation (or, in turn, moral behavior) we are likely to do under them.

¹³ She does discuss the traffic experiments of the Shared Space Movement, but she admits that it is only "partly illustrative." Shiffrin, *supra* note 1, at 1219–22.

¹⁴ See Shiffrin, *supra* note 1, at 1214 (referring to "legal rules" and "legal standards").

¹⁵ Norm typologies can be quite elaborate. See, e.g., Richard T. Morris, *A Typology of Norms*, 21 AM. SOC. REV. 610, 610–13 (1956) (an important example in the field of sociology listing seventeen factors in four categories).

Differences of norm type are only one distinction; there are others. For example, norms, including legal norms, can be distinguished by their norm species. Whereas norm type distinctions arise outside of norm statements, norm species distinctions arise inside of them.¹⁶

Norm statements are made up of at least two conceptual components: operator and content. Norm content is familiar: rules and standards are two types of norm content. A norm's species, however, is determined by the kind of operator inside of it. Operators change the strength of the directive, making it compulsory or not, for example. An important distinction in norm species, and therefore operators, is that between mandatory and aspirational norms.¹⁷ Mandatory norms use operators of strong necessity, such as "must," "cannot," or "have to."¹⁸ Aspirational norms use suggestive or ideal operators, such as "should" or "ought."¹⁹ By putting norm operator and content together, we get a complete norm statement. The directive "employees must wash their hands" is a mandatory rule; whereas the directive "employees should be polite" is an aspirational standard.

While Shiffrin's examples of both legal rules and standards are generally mandatory,²⁰ she does not rule out from her analysis legal aspirational norms.²¹ Below, I will explain how norm species, particularly whether the norms are aspirational or mandatory, can have a dramatic impact on whether legal standards are likely to bring about moral conduct.

I. GOING OUTSIDE THE LEGAL NORM: THE INTERACTION OF NORM TYPES

¹⁶ Brian Sheppard & Fiery Cushman, *Evaluating Norms: An Empirical Analysis of the Relationship between Norm-Content, Operator, and Charitable Behavior*, 63 VAND. L. REV. 55, 68 (2010).

¹⁷ *Id.* at 68–71.

¹⁸ *Id.*

¹⁹ *Id.*

²⁰ Shiffrin explains as follows:

I have in mind standards such as: the unconscionability doctrine in contract law, which deems unenforceable contracts or provisions that are "unconscionable" at the time of formation; the stipulation in contract law that "[e]very contract imposes upon each party a duty of good faith and fair dealing in its performance and its enforcement"; Rule 11's requirement that an attorney conduct a "reasonable" inquiry and attest that pleadings and motions are not presented for an "improper purpose"; the remedial rule providing relief where another has been "unjustly" enriched; and the various requirements in tort and criminal law that one act as a "reasonable" person would under the circumstances or take "due care."

Shiffrin, *supra* note 1, at 1222 (footnotes omitted).

²¹ She expressly allows, for example, for norms that do not have features of mandatory norms. *See id.* at 1231 ("There *might* be serious consequences if a citizen elects not to interpret the legal norm and thereby falls afoul of its requirements, depending on the significance of the norm and the remedies attached to it." (emphasis added)).

When one interprets a legal norm to which one is subject, there are ordinarily two possibilities: either the legal norm will conflict with the nonlegal norms to which she is also subject, or it will not. The following analysis will show how Shiffrin's claims are more likely to be right under circumstances where there is no norm type conflict than where there is norm type conflict. In short, conflict can make legal rules more likely than standards to bring about moral deliberation and moral behavior.

A. The Potential Importance of Norm Type Conflict

It is often the case in a legal system that there is no perceived conflict between the nonlegal and legal norms to which an individual is subject. Sometimes there is no conflict because the person has not internalized many nonlegal norms or those that she has internalized do not appear to her to be particularly relevant under the circumstances. Other times — and this is probably more desirable — the directives of the legal and nonlegal norms appear to her to be in perfect alignment. Under these scenarios of non-conflict, there is a persuasive case that standards are more likely to induce moral deliberation than are rules.

Imagine a storeowner who takes the norms of her religion and of her legal system quite seriously. Further imagine that she is subject to two legal rules. The first requires that the front door to her store push outward toward the sidewalk. Her preexisting religious, moral, or other norms appear to her to be silent on the matter. Since the rule is stated with particularity and there is no conflict with her nonlegal norms, the way is clear for her to interpret and follow the law without moral deliberation. The second law prohibits stores from being open on Sundays and during nighttime hours. It just so happens that her religion is quite clear that Sundays are to remain holy days of observance and that people are required to rest between sundown and sunrise, so all commerce during those days and hours is strictly prohibited. Since she already believed that her religion mandated that she keep her store closed during the legally forbidden times, there is again no conflict and, therefore, no need for moral deliberation in interpreting the law.

Now imagine that she is subject to two similar legal standards. The first requires that the turning function of store doors be installed so that they function "reasonably." She might engage in shortcut thinking. For example, she might choose the swing direction based upon what she has seen elsewhere or been told to do by other shopkeepers. Or she might decide simply based on the direction the screw holes faced when she dragged the door over to the entrance. Regardless, Shiffrin is right that she might consider and balance the public or moral reasons in favor of and against inward-turning doors as the law directs her to do.

Similarly, imagine that the second law requires that stores must keep hours that are in accordance with “the strictest moral standards.” Again, there is a considerable chance that she might ponder what hours of operation would most embody the good. It is also possible, however, that, because she had already internalized norms that provided clear directives on just that issue, she would take a shortcut and quickly access her already-held beliefs and plug them in as the meaning of the standard. Still, the chance that she will engage in moral deliberation seems stronger under standards than under rules in these circumstances.

It is important to mention these various shortcut modes of interpretation because their availability weakens the likelihood that subjects will engage in moral deliberation. Indeed, Shiffrin is well aware of this possibility and offers convincing arguments as to why this risk is not overwhelming.²² As I describe below, however, the risk is considerably higher under circumstances of norm conflict.

Turning now to conflict, imagine that our interpreter feels strongly that her religious norms command that prayer should be the only way to treat her children’s medical issues, but there is a legal rule mandating that all children be inoculated. Because she cares deeply about both norm types, she will feel compelled carefully to make a profoundly difficult decision. She must weigh the importance of fidelity to law against the importance of fidelity to religion. This is arguably legally-induced moral deliberation of the most challenging and intense kind. Moreover, it is entirely consistent with Shiffrin’s criteria for moral deliberation. It neither dictates that “some very specific content be uttered by or thought by citizens” nor “compromise[s] the dignity, autonomy, or privacy of citizens.”²³ Further, it relates “to developing the character traits strongly associated with democratic citizenship”²⁴ because it is an opportunity for an interpreter to strike a balance between liberal rule of law values and personal notions of how to improve the whole of society.

²² *Id.* at 1227–29. More specifically, the short cut in which our interpreter resorts to looking or asking around shares the same defect as Shiffrin’s account of “Holmesian” interpretation; she explains that “[a]lthough I have been arguing that such situations bode well for citizens and the polity because they may induce a salutary form of moral deliberation, dialogue, and sometimes further self-determination, the objector may respond that citizens will evade this deliberation and engage instead merely in sociological prediction, which will not stimulate the relevant forms of moral probing and understanding.” *Id.* at 1228. Whereas the example in which she quickly substitutes her religious belief resembles Shiffrin’s account of displacement thinking, “the posture of the law, when it deploys standards, is not to encourage citizens to displace the law with their own invention. The standard does not serve as a disposable default rule.” *Id.* at 1231.

²³ *Id.* at 1233.

²⁴ *Id.*

Now imagine that the law instead sets forth a standard regarding child health care, such as “parents must act in the best interests of their children with respect to their children’s health care.” It is difficult to imagine how such a legal norm would ever create a conflict that would force the interpreter into intense moral deliberation. The interpreter could see very quickly that some interpretations would conflict with her nonlegal norms, so she would engage in rapid motivated reasoning to choose an interpretation that does not conflict. Since the standard easily permits nonconflicting interpretation, the interpreter is almost sure to succeed with minimal effort. It is possible, of course, that the standard might produce the softer sort of deliberation we see under scenarios of non-conflict, where the person tries to fit various moral options within the umbrella of an evaluative term, but for this behavior to bring about intense deliberation, the interpreter would have to conclude that the precise option she has chosen for the umbrella is at odds with one of her closely held nonlegal norms — very unlikely. It seems much more likely that she will simply engage in the shortcut thinking that we have discussed already.

In this thought experiment at least, it seems that rules would be better at inducing moral deliberation in situations of conflict, and standards would be better in situations of non-conflict.

While Shiffrin does not discuss the role of norm conflict at length, she appears to contemplate that the benefits of rules and standards, respectively, might flip when there are differences between the already-held moral views of the citizenry and the polity:

Pockets of cultural divide or jolts of cultural progress may also mark spots where rules are especially, albeit temporarily, appropriate. Where communities differ with respect to their normative assessments or where the law plays a leadership role in establishing new standards and moral progress, citizens may require quite specific guidance in order to orient or to recalibrate their judgments.²⁵

Importantly, however, she does not appear to concede that inducing moral deliberation is among the things that rules can do better under such circumstances, nor does she appear to think that these circumstances weaken the case for legal standards in any important way.

Admittedly, the persuasiveness of my claim rests heavily on my assertion that it is very unlikely that interpreters would find a way to read legal standards as strong constraints that conflict with their prevailing ideology, thereby forcing them to engage in intense moral deliberation. Fortunately, there is empirical evidence to back me up.

B. The Empirical Evidence

²⁵ *Id.* at 1244–45.

It is not entirely clear what moral deliberation looks like to an outside observer. It probably looks like any other brand of thinking. In a controlled environment, however, it might be possible to get people systematically to describe their thinking process, or to indicate the changing status of their moral beliefs over a period of time, or to lie stationary in an fMRI tube while they solve moral problems, all of which might yield something close to direct evidence of moral deliberation. It is probably easier to gain indirect evidence of moral deliberation. We can expect moral deliberation to produce certain observable effects, and so the absence of these observable effects (or the presence of the opposite effects) under the right conditions can serve as evidence, indirectly, that moral deliberation never occurred. Shiffrin's moral deliberation makes people more likely to engage in moral conduct and, at the very least, makes them more likely to change the content or strength of their moral beliefs.²⁶ Thus, the absence of these features can serve as indirect evidence that no moral deliberation occurred.

Many studies ask people to choose between moral and immoral conduct. If such a study provides evidence that people behave less morally when they are subject to a standard than when they are subject to a rule, then that study also provides evidence that the subjects did not engage in moral deliberation under the standard.²⁷ Similarly, if a study shows that people's moral beliefs are unchanged after being made subject to a standard, then that study indirectly supports the notion that those people did not engage in moral deliberation.

As of yet, there are no studies that have been specifically designed to provide direct evidence of moral deliberation under either rules or standards. Fortunately, there are studies that provide indirect evidence. And while these studies are few in number, those that exist lend support to the notion that standards fail to induce moral deliberation in circumstances of conflict.

Professors Yuval Feldman and Alon Harel used behavioral experimentation to analyze whether rules or standards were more effective at preventing people from following self-interested (and arguably immoral)

²⁶ It is possible that moral deliberation would not change one's moral views at all, but that would be a very unexpected result. If that were expected, then it would be strange, indeed, for Shiffrin to care about its increase at all.

²⁷ It is possible that people could engage in moral deliberation and decide to do worse conduct than they would have chosen if they had not done so. Of course, we would never expect this to be the case, and as discussed, *see supra* note 12, that is not what Shiffrin expects. I should also note that Shiffrin does not appear to claim that moral deliberation is necessary before conduct can be deemed moral conduct.

al) social norms that conflicted with law.²⁸ They asked subjects to pretend that they had changed jobs and that they had the opportunity to convince the largest client of their former employer to continue doing business with them at their new employer. The experiment had four conditions, two having to do with legal norm types (rules and standards) and two having to do with what other people tend to do under those circumstances (one in favor of taking the client and one against). The rule condition clearly forbade convincing the client to come along; whereas the standard condition forbade them from using “unfair practices when competing against your former employer.” As for the social convention conditions, the subjects were told that 78% of people with their job in a survey had reported that, under similar circumstances, either they would take the clients with them (the selfish condition) or they would not take the clients with them (the unselfish condition). Obviously, the selfish condition conflicted with the legal norm. They found that the selfish condition was the most influential of the two.²⁹ More importantly for our purposes, they found that when the selfish condition was paired with a standard, the subjects exhibited the highest overall level of selfishness.³⁰

This result prompted the authors to perform a second study in which they tested how varying levels of reward for selfish behavior interacted with rules and standards. They used the same rule and standard conditions as before but told subjects either that convincing the client would be “significantly profitable” (high self-interest condition) or would “not be significantly profitable” (low self-interest condition).³¹ Similarly to the first study, the results under the standard were the most selfish, and the influence of the selfishness tracked the intensity of the selfish condition; the most selfish results occurred under the combination of the standard and the high self-interest condition, and the second most selfish results occurred under the combination of the standard and the low self-interest condition.³² Thus, the authors concluded that “[s]tandards give people the opportunity to interpret reality in a way that supports their self-interest and hence both noncompliance norms (most people would convince the client) and high gains (if you convince the client you will earn a lot) exert a greater effect when people are faced with [standards].”³³

²⁸ See Yuval Feldman & Alon Harel, *Social Norms, Self-Interest and Ambiguity of Legal Norms: An Experimental Analysis of the Rule vs. Standard Dilemma*, 4 REV. L. & ECON. 81 (2008).

²⁹ See *id.* at 94.

³⁰ *Id.*

³¹ *Id.* at 101.

³² *Id.* at 102–03.

³³ *Id.* at 105.

On the surface, these results might not appear to say much about the power of standards to cause subjects to engage in moral deliberation. The law professors did not set out to test for moral deliberation after all; rather, they tested whether subjects would choose a selfish course of conduct or not. Nevertheless, their results provide indirect evidence that shortcut thinking rather than moral deliberation occurred under the standard. For example, both the first and second experiments showed that those under standards generally choose more selfish options than those under rules.³⁴ To be sure, it is possible that those subject to standards engaged in moral deliberation and, by and large, made more selfish decisions, but even if such deliberation occurred, it arguably does even more harm to the notion that standards are desirable than does the authors' explanation. Moreover, the second experiment showed that the subjects' selfishness under the standard aligned with the level profitability. This result suggests that the subjects simply considered the intensity of their selfish motives and did not engage in more complex analysis such as moral deliberation.

Professors Adam Cox and Thomas Miles studied a similar phenomenon, but rather than using behavioral experimentation they used statistical methods to analyze actual court decisions under the Voting Rights Act.³⁵ They chose to focus on the Act because a 1986 Supreme Court decision adopted a two-step test for applying the Act, thereby creating a valuable jurisprudence for the study of rules and standards: in particular, the Court adopted a rule-like doctrinal test and a standard-like totality of the circumstances test.³⁶

Cox and Miles discovered that after the change in the Act's interpretation, "[i]deological divisions in judicial voting patterns [became] more pronounced in the standard-like second step . . . than in the evaluation of the more rule-like factors . . ."³⁷ In particular, between 1986 and 1994, Democratic appointees were much more likely to find a violation than Republican appointees (91% compared to 59%), a difference in rates that was twice as high as the difference under the rule prong for each party.³⁸ Here again we see norm conflict — namely, conflict between the ideological norms for Democrats and the legal rule. And the standard allowed the judges to reach decisions that were consistent with their (probable) already-held beliefs. That is, the de-

³⁴ *Id.* at 94, 102–03.

³⁵ Adam B. Cox & Thomas J. Miles, *Judicial Ideology and the Transformation of Voting Rights Jurisprudence*, 75 U. CHI. L. REV. 1493, 1494–95 (2008).

³⁶ See *Thornburg v. Gingles*, 478 U.S. 30, 48–51 (1986).

³⁷ See Cox & Miles, *supra* note 35, at 1495.

³⁸ See *id.* at 1524. To be sure, this effect was less pronounced after 1994, but the authors partially attribute this differential to changed partisan and representational consequences of vote dilution litigation after that date that would lead to the Democrats viewing the tool of the Voting Rights Act less favorably. See *id.* at 1503–05.

cline does not appear to be the result of a moral deliberation that, over time, shaped the initially Democratic-leaning judges into moderate judges. The ideology came first, not the standard.

It might be objected that these studies do not provide a clear measure of moral deliberation. My own recent empirical research fills this void somewhat. In a forthcoming piece, I analyze the power of rules and standards to constrain judges in a context with the same legal/nonlegal norm conflict I described above.³⁹ In that study, I further tested rules and standards to determine first, which one requires more time for those applying it to reach a convincing result, and second, which one is more likely to cause a change in the moral feelings of those applying it. It is this last measure that serves as an indicator of moral deliberation. A brief description of my methodology explains how.

I asked subjects to act as judges⁴⁰ in a simulated, ideologically divisive asylum case.⁴¹ In a nutshell, they were asked to decide whether to grant asylum to a politically persecuted immigrant who had overstayed his visa.⁴² In one phase of the experiment, subjects were able to decide whether to grant asylum without the assistance of any law on the issue.⁴³ In a second phase, subjects received the same fact pattern except with the addition of either a rule or a standard that related to whether asylum ought to be granted.⁴⁴ The rule quite clearly forbade the immigrant from applying for asylum because he had waited too long.⁴⁵ The standard stated only that the immigrant must file for asylum within a reasonable amount of time.⁴⁶ In each phase, subjects were told to write an opinion justifying their decision and to state the strength of their conviction that they had done the right thing on a 1-10 Likert scale.⁴⁷ For approximately half of the subjects, I gave the legal interpreters as much time as they needed to write what they felt

³⁹ Brian Sheppard, Capturing Constraint (unpublished manuscript) (on file with the Harvard Law School Library).

⁴⁰ It might be objected that my study is less relevant to Shiffrin's essay because it asked law students to pretend that they were judges whereas Shiffrin's essay focuses on deliberation from the "citizen's perspective" and not on how "standards in constitutional law induce adjudicators to deliberate and to render decisions that make the adjudicator accountable and transparently responsible." Shiffrin, *supra* note 1, at 1217. While there are certainly differences between my subjects and actual judges, it should be noted that my subjects shared important features with actual judges, most notably that they were not making publicly transparent decisions that held them personally accountable. Indeed, they were guaranteed anonymity pursuant to the recommendation of a human subjects committee.

⁴¹ Sheppard, *supra* note 39, at 40.

⁴² *Id.* at 41.

⁴³ *Id.* at 41-42.

⁴⁴ *Id.* at 42-43.

⁴⁵ *Id.* at 43.

⁴⁶ *Id.*

⁴⁷ *Id.* at 41.

was a convincing opinion, and on the other half, I strictly limited the amount of time that they had.⁴⁸ Thus, there were two independent variables — norm type (rule or standard) and time limitation (time unlimited or time limited) — and two dependent variables — outcome and strength of moral conviction.⁴⁹

The experiment provides at least three indicators of moral deliberation: first, whether a legal interpreter's moral convictions changed after the intervention of a legal norm; second, whether the time available for deliberation had an effect on those convictions; and, of course, third, whether the introduction of the legal norm changed the outcome of the case. These indicators provide insight into whether moral deliberation occurred by measuring both the change in the moral conviction and in the relevant decisions of the subject with respect to an identical set of facts under relevant law (either a rule or standard) and under no relevant law. To be sure, it does not allow us to enter into the minds of the judges, but it does allow us to see whether there has been a change in their moral beliefs and in the decisions they reach.⁵⁰ Furthermore, moral deliberation takes time, so we can assume that there would be an interaction between a change in the strength of the moral conviction and available time. In particular, when one has the time to engage in moral deliberation during legal interpretation, it is likely to increase the confidence in the interpreter that the subsequent, chosen interpretation represents the right thing to do, whereas strict limitations on time will not be likely to produce the same increase. In addition, we can assume that law-induced moral deliberation will be likely to produce a significant difference in case outcomes. It is less likely that moral deliberation has occurred where the legal outcome is the same regardless of whether there is legal norm on point in the case. This is not to suggest that the difference ought to be equal as between rules and standards — we can expect rules to outperform standards on this score — but if there is substantial moral deliberation happening under a standard, there ought to be a statistically significant difference in case outcomes as a result of the imposition of a legal standard.

The results support finding that the standard did not increase moral deliberation. Indeed, the behavior of those subject to rules, as opposed to standards, was consistent with the expected behavior of those that engage in moral deliberation. Beginning with the most basic indicator, outcome, rules brought about a statistically significant change in outcomes, whereas standards did not. Under rules, 26.8% of subjects changed positions from their non-rule determination.⁵¹ Only 6.5% of

⁴⁸ *Id.* at 43–44.

⁴⁹ *Id.* at 48.

⁵⁰ See *supra* notes 26–33 and accompanying text.

⁵¹ Sheppard, *supra* note 39, at 58.

those subject to standards changed positions.⁵² In addition, those subject to rules exhibited greater changes in the strength of their moral conviction, and in precisely the directions we would imagine under moral deliberation. Those subject to rules that were subject to strict time limits *dropped* 1.4 points in their moral conviction, and those subject to no time limits *increased* .4 points.⁵³ Contrariwise, those subject to standards under strict time limits *dropped* by only .2 points and those subject to no time limits *increased* only .1 points.⁵⁴ The change in conviction for rules was significant, but the change in conviction for standards was not.⁵⁵ So unlike rules, standards did not display an interaction with available time; giving subjects the opportunity to take as much time as they wanted did not lead to a significant difference in their moral conviction. Thus, by allowing people to have an opportunity for moral deliberation, there is every indication that those subject to rules, and not to standards, used the opportunity and ended up strengthening their sense of moral justification in their result. This evidence supports the notion that those under standards did not engage in moral deliberation.

These differences become even more pronounced when we look only at those subjects that felt the most constraint by the legal norm — namely, those who switched positions after interpreting the law. For outcome flippers subject to a rule, those under the time limit saw a colossal *drop* in their moral conviction of 4.3 points; whereas those who were subject to no time limit saw a substantial *increase* in their moral conviction of 2.2 points.⁵⁶ Comparing this to outcome flippers under standards, there were absolutely no flippers under the standard with no time limit, but the overall average for flippers under the standard was only -.09.⁵⁷ Of course, because there were so few flippers under the standard, these numbers are not statistically significant.

To be sure, the findings do not rule out the possibility that those subjected to standards were engaging in moral deliberation that led neither to a significant change in outcomes nor in moral conviction, but it is much more likely that moral deliberation was occurring under rules than under standards.

Thus, the existing empirical data lend more support to the account of standards under norm type conflict than to Shiffrin's account: legal interpreters that are ideologically motivated generally appear to treat legal standards as little more than a legal blessing to tap into their pre-

⁵² *Id.*

⁵³ *Id.* at 82.

⁵⁴ *Id.*

⁵⁵ *Id.* at 54.

⁵⁶ *Id.* at 64.

⁵⁷ *Id.* at 83.

existing moral beliefs, allowing them quickly to reach the result that they favored before they bothered interpreting the legal standard in the first place.

As a result, it might be wise to limit Shiffrin's case for standards to those situations in which there is no conflict between legal and nonlegal norms.

II. GOING INSIDE THE LEGAL NORM: THE INTERACTION OF NORM SPECIES

In discussing the interaction of norm *species*, my focus will be on a situation in which we know what the most moral conduct is — namely, charitable donation. The moral status of charitable giving shouldn't be in serious doubt; this conduct is understood by most to be supererogatory.⁵⁸ I will show how standards only succeed in maximizing the moral conduct of charitable donation when they are housed within a particular, and less popular, legal norm species. Otherwise, they prove to be worse than rules.

Recall that the operator of a norm statement tells us its species. Important here are two species: mandatory and aspirational norms. While Shiffrin uses examples of mandatory legal norms when she discusses the legal standards she has in mind, she fails to address aspirational legal norms. This could be a mistake. Aspirational norms are a valid and commonly used legislative tool.⁵⁹ It might be fine for Shiffrin to omit aspirational norms and other norm species from her analytic framework if the interaction between rules or standards and the kinds of operators to which they are attached were insignificant. The empirical research of this issue, however, suggests the opposite.

Recently, Fiery Cushman and I published the results of an experiment in which we gave law students an opportunity to donate money to a legal services charity.⁶⁰ Specifically, our subjects were asked to

⁵⁸ See, e.g., David Heyd, *Supererogation*, THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Fall 2008 Edition), <http://plato.stanford.edu/archives/fall2008/entries/supererogation/> ("Beneficence and charity are often considered as typical examples of supererogation."). But cf. Peter Singer, *Famine, Affluence, and Morality*, 1 PHIL. AND PUB. AFF. 229 (1972) (arguing that certain conduct commonly understood to be charitable is subject to a moral duty).

⁵⁹ The most familiar examples of aspirational legislation are probably those drawn from ethics codes, environmental law, or international law, but at least one scholar has argued that we can interpret some provisions of our Constitution as having an aspirational character. See Richard H. Fallon, Jr., *Judicially Manageable Standards and Constitutional Meaning*, 119 HARV. L. REV. 1274, 1324–25 (2006) ("If a constitutionally permissible gap can exist between background rights and the rights that courts or other officials will currently enforce, the best rationalizing explanation may be that background rights can be (though not all of them need to be) partly aspirational, embodying ideals that do not command complete and immediate enforcement." (footnote omitted)).

⁶⁰ See Sheppard & Cushman, *supra* note 16, at 74.

play a modified version of The Dictator Game, in which we gave them \$10 and then subjected them to one of six possible norms, each having to do with the appropriate division of the money between them and the charity.⁶¹ Each norm represented a unique combination of a mandatory or aspirational operator and either a rule or a standard, paraphrased as follows:

(1) Aspirational Standard: You ought to give according to the strictest moral standards.

(2) Weak Aspirational Rule: You ought to give at least \$3.

(3) Strong Aspirational Rule: You ought to give at least \$7.

(4) Mandatory Standard: You are required to give according to the strictest moral standards, and you could be punished for noncompliance.

(5) Weak Mandatory Rule: You are required to give at least \$3, and you could be punished for noncompliance.

(6) Strong Mandatory Rule: You are required to give at least \$7, and you could be punished for noncompliance.⁶²

The worst performer of these various norms was the mandatory standard. The mean level of giving was only \$4.56; whereas the mean level of giving under the aspirational standard was \$6.17.⁶³ This predicted effect was statistically significant.⁶⁴

Mandatory standards also compare unfavorably to mandatory rules. Both weak and strong mandatory rules outperformed the mandatory standard, leading to mean giving of \$5.03 under the weak rule and \$6.85 under the strong rule.⁶⁵ The differences between the mandatory standard and both the weak and strong mandatory rules were significant.⁶⁶

Things were not all bad for the standard, however. Aspirational standards happened to spur more giving than did the weak mandatory

⁶¹ *Id.* at 71-74, 77-78.

⁶² *Id.* at 77-78.

⁶³ *Id.* at 81.

⁶⁴ *Id.* The means tell only part of the story. When aspirational norms produced higher rates of giving, the result was driven principally by a higher proportion of subjects giving exactly \$10, and a lower proportion of subjects giving exactly \$0. *Id.* at 81-82.

⁶⁵ *Id.* at 81.

⁶⁶ *Id.* at 82.

rule (\$5.03) or the weak aspirational rule (\$6.05), but these differences were not statistically significant.⁶⁷

Why did this occur? The best theory is that mandatory norms have the capacity to crowd out intrinsic motivations to behave morally. That is, when one has an intrinsic motivation to behave charitably, for example, the intervention of a reward or penalty for doing or refraining from moral conduct can make the person lose his or her intrinsic motivation and become a selfishly motivated gamesman with respect to that intervention.⁶⁸ Moral deliberation this is not. To riff on a famous illustrative example,⁶⁹ the intervention of, say, a \$5 reward for blood donation can reduce the overall amount of blood donation from a rewardless system because those that otherwise would have given blood for free are now weighing whether their blood is worth \$5.⁷⁰ If that amount is found wanting by a high enough proportion of the population, the reward can lead to an overall reduction in blood donation.

The good news about motivation crowding in the charitable giving context is that it appeared only to occur under mandatory norms. The bad news is that it appeared to occur regardless of whether the norm was a rule or a standard. And this is particularly bad for standards because they lack the rule's power to overcome crowding through strong incentives to perform clear minimum conduct. In other words, only rules have the capacity to curb the selfish behavior created by the mandatory operator.

The lesson is that mandatory standards are a risky means of bringing about an increase in a moral conduct when those subject to them are already motivated to perform that moral conduct. As discussed, the capacity of legal standards to bring about an increase of moral conduct is arguably even more important than their ability to bring about moral deliberation, so the finding that rules can prompt more charitable donation than standards, even when they are not mandatory, is troubling.

CONCLUSION

⁶⁷ *Id.* at 81–82.

⁶⁸ *Id.* at 63–64.

⁶⁹ RICHARD M. TITMUSS, *THE GIFT RELATIONSHIP: FROM HUMAN BLOOD TO SOCIAL POLICY* 245–46 (1971).

⁷⁰ Shiffrin discusses a similar phenomenon when she describes the Shared Space Movement and its success stories in European traffic. See Shiffrin, *supra*, note 1 at 1219–21. Her description of how traffic rules can crowd out a preexisting motivation to drive safely bears some resemblance to my description of motivation crowding: “I draw the lesson that the background standard that one is to ‘drive safely’ may be more thoughtfully deployed without a myriad of specific signals and rules that, in some contexts, spark complacency and automatic behavior.” *Id.* at 1220–21.

I hope that this analysis shows how norm type and norm species can influence the capacity of legal standards to bring about moral deliberation and behavior. Before finishing, it is worth considering what the *combination* of these two norm dimensions might say about this capacity. With the caveat that further study is necessary, it could be that for standards to operate in the way Shiffrin anticipates, it is generally necessary that those subject to a standard do not detect something to which they are averse and which appears to be linked to the standard. The interpreter might be averse to being commanded to do something moral when she already would have done something moral in the absence of the command. Or it might be that she detects that interpreting the standard in the thorough way Shiffrin desires could bring about a serious conflict between a legal norm and another norm that she also holds dear. Or it might be that she is hurried, and interpretation of a standard seems like too much effort under the circumstances. In contrast, Shiffrin's case for standards appears to be strongest when those subject to them are at peace, without such worries.

Despite these speculations, there can be little doubt that Shiffrin's essay is a valuable contribution to the rules versus standards debate. In addition, her elaboration of the relationship between legal norms and moral deliberation, as well as her description of moral deliberation, ought to serve as the conceptual spadework for those seeking to study empirically how legal standards influence thinking and behavior.

I am confident, however, that even the small amount of existing empirical scholarship on this subject illustrates that a simple legal rule/legal standard distinction is inadequate to understand how norms bring about moral deliberation. Furthermore, it is already obvious that commonplace factors extrinsic to law — such as the motivation to follow nonlegal norms, especially selfish ones — deserve consideration as they appear to be able to flip the relative strengths and weaknesses of rules and standards discussed in this Response.

Despite their shortcomings, standards do not deserve to be thought of as second-rate norm content. This Response does nothing to weaken the claim that standards can serve as the impetus for moral deliberation when there is neither a serious conflict between legal norms and nonlegal norms nor strong, conflicting selfish motivations. When such conflicts arise, however, the standard is something of a weakling, serving as only a mild constraint on the nonlegal motivation. Under the same circumstances, the rule stands up for itself, often forcing the interpreter to engage in what appears to be intense moral deliberation. But there is no reason to assume that legal standards will be so ineffectual when they are paired with mandatory norms in a setting that is not supererogatory, especially one where the appropriate moral conduct is not so clear.

It might further be asked, then, whether these two scenarios — when the law regulates clear supererogatory conduct and when the law is in conflict with serious nonlegal motivations — are so commonplace that it is not worth throwing our support behind Shiffrin’s claim. I would suggest that the answer is no. American law only rarely regulates supererogatory conduct,⁷¹ and rarer still does it mandate such conduct.⁷² Also, the majority of decisions that people make within a legal system do not involve serious conflict between the legal directive and the decider’s personal motivations as described (otherwise our legal system would be in deep trouble!). Nevertheless, the existing empirical scholarship on this issue has thrown salt on some of standards’ supposed virtues. It could be that additional study will only operate to weaken them further.

⁷¹ See, e.g., 26 U.S.C. § 170 (2006) (setting forth tax deduction for charitable contribution).

⁷² See Ken Levy, *Killing, Letting Die, and the Case for Mildly Punishing Bad Samaritanism*, 44 GA. L. REV. 607, 616 (2010) (identifying only four examples of laws in the United States that mandate Good Samaritan conduct).